On the Horns of a Dichotomous Dilemma

by

Jay C. Powell, James Bernauer and Vishnu Agnihotri

Heisenberg's *Uncertainty Principle*[1] makes the point that the way we observe predetermines the conclusions we can draw. Presenting the universe as a dichotomy of being either digital or analog but not both may be playing into this same dilemma. This essay takes the same position as Heisenberg; reality is *both* analog and digital. This paper will illustrate this perspective by presenting our research into educational assessment.

Current test-scoring practices reduce performance into digital terms in the form of the frequency of acceptable answers on tests. This produces a single number that is comparable statistically with other numbers from the same data set using procedures derived from the *general linear model* (GLM) as it relates to classical test theory.[2] This theory presumes that alternative answers contain no useful information about performance. Alternative answers can be reduced, therefore, to zero (0); making each response binary and the data set reducible to values from the whole number set N_0. This reduction produces an artificial dichotomy that allows the use of GLM procedures. This essay challenges the validity of this digital reduction of learning behavior.

It has been long known that there is a possibility that the Gaussian curve, the "normal distribution," represents random events very well. On the other hand, it may conceal complicated or complex substrata, if the events being observed are not random.[3] This perspective is particularly true when linearity is incorrectly presumed. At this point procedure may lead to a destructive over-simplification of these data.

Are Wrong Answers Educationally Meaningless?

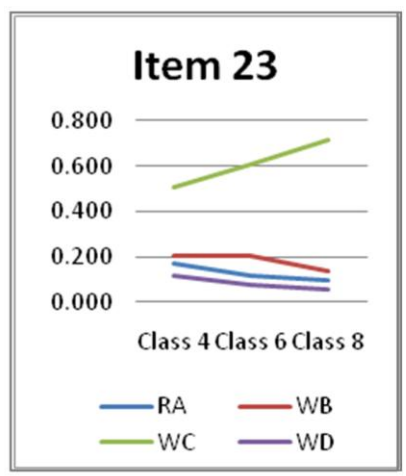
Let us consider a case in point. Here is a question from India[4] given to about 16,000 students distributed unevenly at years 4, 6 and 8.

Figure 1: A Sample Multiple-Choice Test Item

23 The night of July 31st, 2004 is a full moon in Delhi. Will it also be a full moon in Mumbai? In New York?

- **A.** It will be a full moon night in both Mumbai and New York*
- **B.** It will not be a full moon night in either Mumbai or New York
- **C.** It will be a full moon night in Mumbai, but not in New York
- **D.** It will be a full moon night in New York, but not in Mumbai

Item 23	Class 4	Class 6	Class 8
R_A	0.168	0.114	0.094
W_B	0.206	0.202	0.137
W^C	0.509	0.608	0.713
W_D	0.116	0.077	0.056
Totals	1.000	1.000	1.000



Almost no one got this question right. Their results began at about one in eight and dropped by half by Class 8.

The wrong alternative C increased by about 10% at each level from one in two in Class 4 to almost three in four in Class 8. This is a statistically significant proportional gain (z = 2.9).

Presumably, they do not understand the spatial relationships in our solar system. If teachers explained this concept to them, they did not understand the explanation. Evidently, they have learned a misconception. It is true that New York is opposite to Delhi on earth. They have generalized this relationship to the solar system. Such scientific misconceptions led to the development of concept inventories[5].

It is also interesting that one in five chose explanation (B) that is contrary to the facts.

The third option, alternative D is the opposite of C.

These students have been taught a truth about the relationship of their country to the eastern United States on the earth. Generalization is a legitimate thinking strategy.

This answer is obviously *not* a random event. It is a learned misconception. When we disregard this answer by scoring it zero (0), we lose the ability to identify and correct such errors, either specifically or in general for all cultural over-generalizations and misconceptions. The conclusions they have drawn reflect the ways they have interpreted the question.

Our research has shown repeatedly that most alternative answer selections come from thoughtful interpretation of the questions, as just illustrated and are *not random*.[6]

Augmenting Observations with Repeated Measures

If the alternative answers are not random, perhaps giving the same test twice to the same group will give us a picture of the dynamics of learning

Consider the following example. This item, from Gorham's (1957) *Proverbs Test.*[7] It was given in two administrations (October and the following March) to 2,810 students from the third grade through the end of high school.

The group reported here consisted of those students who gave complete answers to this and another instrument (Schools I Would Like to See, Powell, *et al.,* 1997)[8] on both administrations. The full sample was 3,000+. The age range is indicated along the horizontal axis from < 96 months to > 220 months. We tested all the children in 10 schools (eight elementary and two high schools).

Figure 2: A Sample Test Item Given Twice

QUICKLY COME, QUICKLY GO (Easy come, Easy go).

A) Always coming and going and never satisfied.
B*) What you get easily does not mean much to you.
C) Always do things on time.
D) Most people do as they please and go as they please.

The solid arrows represent a multinomial *Thurs* (*p*) > 0.945.[9] The hollow arrows were less than this but close to it and seemed to fill missing spots in a sequence.

The response interpretations collected for the alternative answers to this question are included in the discussion. Their reasoning reports came from *interviews of a representative sample of students on an earlier administration of the same test with* the indicated age groups.[10]

Alternative C "Always do things on time" had a most common selection (mode) for eight-year olds. Their most common explanation was, "That's what the teacher always says." Picture the frenetic activity of the typical third grade classroom and considering only the "Quickly come, quickly go" part of the proverb as presented and their answer makes sense. This answer is an age-appropriate "egocentric" (in Piaget's terms) response.

Alternative D was most commonly selected by 10-year olds. Try matching "Quickly Come, Quickly Go," with this answer, "Most people do as they please and go as they please." The language similarity is not coincidental. These are concrete thinkers and they are trying to find the closest literal match to their thought processes. Once again, this is a developmentally appropriate answer (Concrete Operations, in Piaget's terms).

With the thirteen-year olds, the most common response was alternative A "Some people do as they please and go as they please." Once again, these students are simplifying the item by paying attention only to the "Quickly …" part and not to the "Easy …" part of the lead statement. In this case, they translated it into "A rolling stone gathers no moss" by substituting this thought for the "Quickly …" phrase. It is clear that this reasoning is more sophisticated than that of the earlier age groups. They are not yet culturally correct in this translation, but the behavior is appropriate for the transition from concrete to abstract thinking. (Piaget did not label ths stage.)

The right answer, of course is alternative B. This selection occurred most commonly with sixteen-year olds. Theses ages fit with Piaget's[11] developmental stages.

We classified this right answer as representing a *narrow perspective* in thinking using the *Schools* … opinion survey. We used *Schools*… in the place of interviews in the 1977 study. We associated *Schools* … (a qualitative data opinion survey), with *Proverbs* … (a quantitative reading comprehension test) using crosstabulation. The *Schools* … test is an analog test that uses answer-patterns scoring to produce six scales from eight responses. This feat is accomplished by scoring the entire test *as a unit* in three different ways. One approach produces four scales (1, 2, 3 and 6).

Our purpose was to distinguish between low and high demand "right" answers. Another study had suggested than giving "right" answers and "understanding the concept" may be unrelated events.[12] Bond and Fox[13] addressed a similar issue by commenting upon students who answered the more difficult items correctly and the easier ones incorrectly.

The proportion of right answers declined after age 16. With the students older than this and with adults, we observed that a shift away from the right answer to this wrong one occurred among a meaningful minority[14] of examinees. In this case, the adult reasoning included the "Easy …" part, recognized the "right" answer, and then explained that for the pejorative use of this proverb, "never satisfied" was a better match. By bringing sociological considerations into their answer, they gave a valid answer that was not anticipated by the designer of the test. We suggest that this answer exceeds the ceiling of the test, identifying profoundly informed students and, in this case, making this answer correct as well.

This observation is one of the most important discoveries of this research. It means that we can recognize when people move out of "Formal Operations" into a more elaborate (multi-dimensional logic) form of thinking. Being able to identify when this cognitive transformation occurs empowers us to determine the conditions within which this change into metacognition is teachable.

The Linear Dependency Problem

Having crosstabulated every answer against every other answer in this repeated measures setting, the need was to bypass the linear dependency problem inherent in the use of every answer when applying GLM statistical procedures to test data. Linear dependency arises because the sum of all the right and all the wrong answers is a constant, the number of test items. Powell and Shklov (1992)[15] develop a procedure that bypasses this problem.

The equation for this procedure is:

$$\text{Þ} = P_O / P_T \qquad (1)$$

Where P_O is the cumulative probability of the observed frequency or less in each cell and P_T is the cumulative probability of the maximum possible frequency in that same cell.

The main equation is:

$$P_O = \sum_{k=O_{(min)}}^{O_{ij}} \frac{N!}{o_k!\ s_k!\ t_k!\ u_k!} (p_1)_k^o (p_2)_k^s (p_3)_k^t (p_4)_k^u \qquad (2)$$

The Thurs (þ) statistic is a new parameter. This symbol is being introduced by the authors to represent the *proportion* of explained variability that the observed frequency or less contributes to the total possible frequency that cell can provide. The procedure is applied to each cell of a crosstabulated event matrix. This symbol is used because both lower- and upper-case *P*s already have statistical meanings as individual and cumulative probabilities. In this case, *P* is a cumulative probability of the observed frequency or less.

There are only four parameters used in the two dimensional case because our concern was to consider each cell in turn as it related to the entire matrix. This approach means that the remainder of the table is collapsed around each cell. The purpose is to retain the structural relationship of that cell with the rest of these data and not to impose assumptions upon these relationships. Instead, we wish to discover what these structural relationships might be. In this way, we bypass the linear dependency inherent in the GLM and access any non-linear substructure that might be present among these data.

Because the Þ is a proportion (range 0.000 to 1.000), it is analog, rendering a *distance* relationship among cells from the digital frequency count. As a result, this procedure recovers analog relationships from digital data. A value of 1.000 indicates spatial equivalence and of 0.000 indicates complete disjunction.

This item was chosen for discussion purposes because it is the only item in the 40 on the test that has five of the six performance categories our research uncovered. The *right answer* on this item is classified as representing a narrow thought perspective. There are 24 such answers, as determined by associating their worldviews (analog) using Scale 5 of the *Schools* … opinion survey. The remaining 16 have a broad perspective on this same basis. There is some correspondence between easy questions and a narrow perspective.

This definition-generating procedure associates digital (observation specific) responses with analog (structured-opinion-based) responses to separated those answers more closely related to memory from those more closely related to insightful understanding. This distinction is not made using scoring that considers only the frequency of acceptable answers.

In the overall structure of this scoring system, we have seven categories of response instead of two (right/wrong). The original author of the test provided 20 concrete right answers along with the 40 abstract right answers. This left 100 alternative answers to be classified.

Our work separated the all right answers into 2 categories, and 12 of the wrong answers into answer into two categories, one in the developmental position below Formal Operations and one above it for profoundly-informed learners.

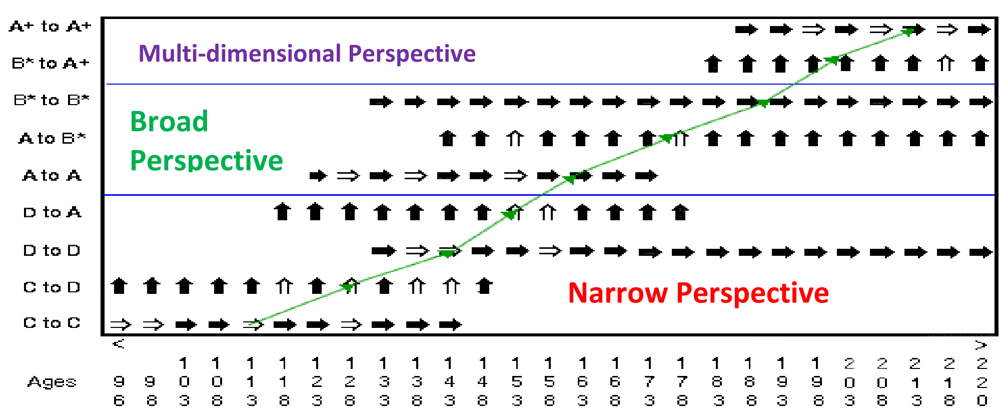
Having shown that there is a reciprocal relationship between reasoning and answer selection with adults,[16] the possibility of a similar relationship with children was hence researchable.[17] This test was administered to 550 children from the third through the eighth grades. Interviewing a representative sample of them by trained interviewers immediately after its administration obtained their reasoning behind their errors. The results of these interviews are the basis for the above reports.

The alternative answers were isolated from both right-answer subsets, clustered and the set into subsets of at least four members. These subset scores were combined along with the two right-answer subset scores and clustered again. With the wrong-answer subsets classified by the age modes of selection, this the emergent order of selection perfectly recapitulated the age order of the modes.

This correlation matrix containing both right and wrong answer subscores has subsequently been shown to be conjoint.[18,19] Mathematically; there is no longer any question that alternative answer selections contain student-performance information. This conclusion has now been supported in six different studies on two continents with people from less than eight years old through adults.

Graphically, these outcomes appear like this:

Figure 3: Item 18 -- *Proverbs Test* –Change Dynamics -- Cognitive Gains

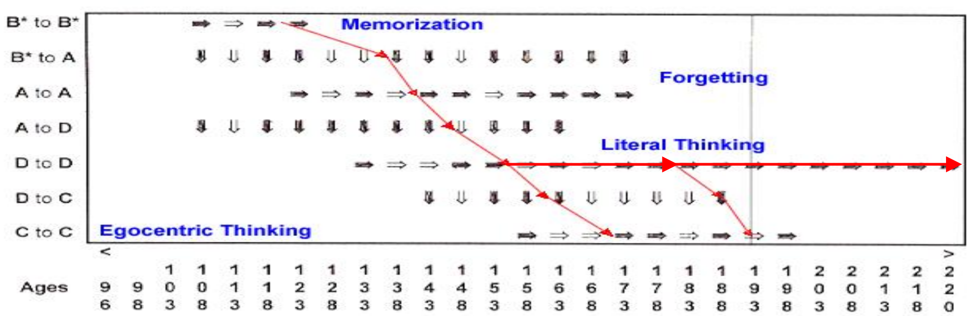


To address this flow, we needed to establish the strength of the relationship in every cell at every age level for every question. This gave a 16 by 27 matrix for each of the 40 items on the test. From here, it was necessary to show that the modes of answer changes occurred appropriately between the modes of the repeated (stable) selections. In this chart, the arrow heads are at the modes. More than 80% of these transformations were appropriately placed on all 40 items. This agreement is sufficient to establish a developmental sequence.

The bold-faced arrows were at or above our arbitrary criterion þ > 0.944. This criterion is arbitrary because we have yet to determine the measurement error of this new statistic. The hollow arrows did not meet this criterion, were usually close to it and there were no more than two of these in a row. Thus, the procedure makes performance information about students available from the "unacceptable" answers selected on this test. The dynamics of learning supporting Piaget's clinical observations emerges. These results are both quantitative (digital) in that they are frequency counts of answer selection pairs, and qualitative (analog) in that their interpretation reflects the thought processes behind their selection.

In this chart, all the transitions are shown ascending the developmental scale. This observation raises the question, is there any a similar pattern for those students whose changes moved in the opposite direction? Our next chart answers this question.

Figure 4: Item 18 -- *Proverbs Test* –Change Dynamics -- Cognitive Losses



This second pathway shows multiple pathways to leaving school early. Most of the people who did not write in March, because they had left school voluntarily, selected either D (literal) or C (Egocentric) thinking in October. Notice the strength of the stream of literal thinking students who finish high school (horizontal extension of the red arrow).

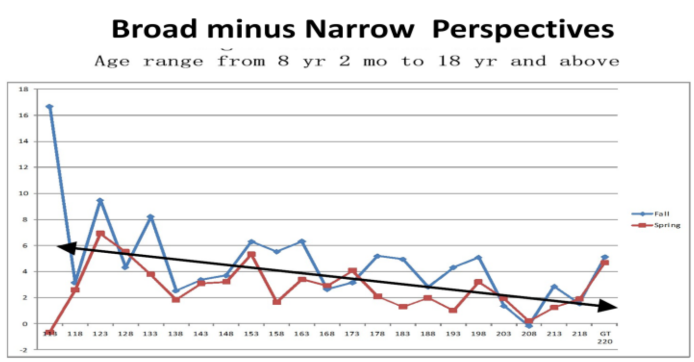
In a more recent study with college undergraduates, we used an exploration-of-ideas approach to teaching to replace the traditional lectures. One class gained in scores with little cognitive improvement, while the other declined in scores, with considerable shifting away from the "right" answers towards the "wrong answer" subtest that reflects increasing cognitive skill beyond the ceiling of the test.

This transition was supported by corresponding shifts in the *Schools* … survey from a narrow perspective in the pretest to a broad perspective in the posttest. A replication of this study may make the reasons for these inconclusive outcomes evident.[20] Additional study could identify the parameter complex that makes this transition with maturing students a general result of teaching.

Another issue is that this item is the only *all-levels* item. All the others have least two of the options at the same level, differentiated by selection strategy. These bifurcations may signal other pathways. A replication of more than one year is needed. The additional year(s) will make it possible to follow students making these bifurcations further to determine whether these alternative routes represent additional pathways or digressions.

Another major observation made from these studies is shown next:

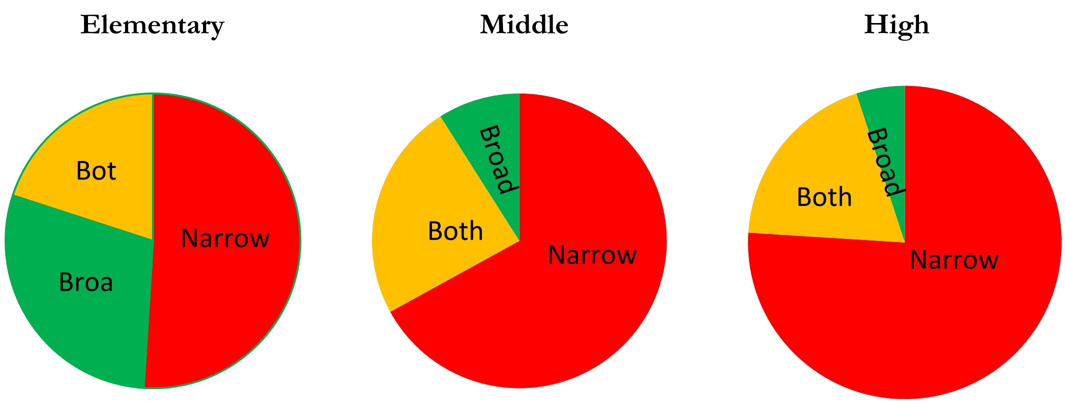
Figure 5: Change dynamics in Perspective (October to March)



In this chart, the red line for March is mostly *below* the blue line for October. This observation indicates a steady *decrease* in overall broadness of perspective. This decrease occurs during the school year with some recovery over the summer. Are we reducing student thought flexibility by the way we educate them? Notice the dramatic drop in the third grade, signaling the fourth grade slump.

The situation is even more dramatic among the top 20% of the scorers. The following pie charts give this picture.

Figure 6: Apparent Impact of Schooling upon High Scoring Students



Taking the top 20% of the students at each level, here is what happened to their thinking perspectives. The broad perspective nearly disappeared, the narrow perspective increased from 50% to 70% and the proportion of those students who were high in both subscores remained relatively stable.

This ability to balance broad and narrow perspectives, contrary to the previous chart showing progressive narrowing of perspective as the average event, suggests that this skill is not learned in school. However, it seems to protect these students from the deleterious impact of information-transmission-based teaching.

A follow up with college undergraduates[21] suggests that the transition to a multi-dimensional perspective has can be taught. In one class, out of two, students moved to higher cognitive skills, acquired this balance between narrowing and broadening perspective. Such an approach uses mind-expanding insights and attention to details in combination to get students to construct their own knowledge.[22] The model represents a conjunction of analog and digital approaches to learning.

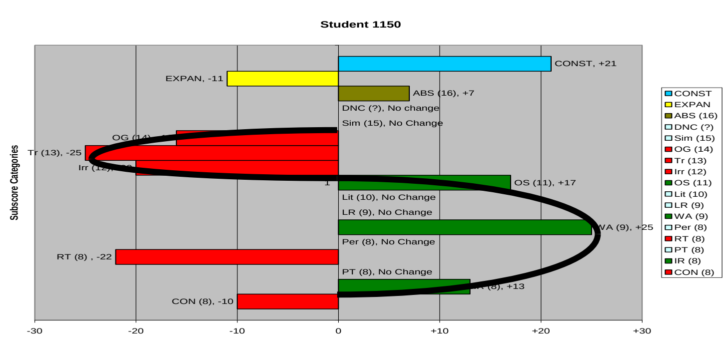
Our opinion survey, *Schools I Would Like to See*, shows a variety of ways to find a balance among the eight approaches to education encoded therein, along with similar balancing of other dimensions in the survey.

Application to Individual Student Learning

There is another way in which learning is analog. The lower level categories of answer diminish and the higher ones emerge with advancing cognitive development. The reverse occurs with declining development. If plotted upon a vertical line representing chronological age, the changes appear as an ascending or descending *sine* curve.

Here is one example (Red bars are declines and green bars are increases in subtest scores):

Figure 7: Increasing Total Score with Declining Cognition



This 13-year-old student has declined in thinking skills from functioning like a 14-year-old (longest bar to the left) to functioning like a 9-year-old (longest bar to the right) in five months. The blue bar at top right represents an increase in narrow perspective by 6 answers, the yellow bar at top left is a decline in broad perspective by 3 answers and the third bar from the top is an increase in total score by 3 answers. (The numbers at the ends of the bars are percentage change.)

A more comprehensive discussion of providing individualized reports from all answers is presented in Powell (2010b). It is sufficient to say that total-correct scores cannot provide as detailed a level of performance information. The standardized gain, using a linear-scale average, is two answers each year. Student 1150, whose performance deteriorated by five years using profile change analysis, would be credited with *a year and a half gain in five months* when using only the total-correct score change is considered! This counter-intuitive observation, demonstrating the invalidity of exclusive consideration of total-correct scores, becomes our main objection to reducing performance to a dichotomy. The current procedure overlooks the analog aspects of learning.

Implications

This paper presents the results of analysis of student responses to a 40-item multiple-choice reading comprehension test using a statistical technique that bypasses the linear dependency problem. The test was given to the same 2,810 students twice to pick up the change dynamics as learning progressed during a five-month interval (October to March).

These are the findings:

1. Considerably more information about each test taker is available from this procedure than can be recovered from the frequency of acceptable answers.
2. The outcomes from total-score changes do not necessarily agree with the actual changes in cognition that is occurring, while the changes in answers to the broad perspective and narrow perspective right answers subscores and the wrong answer sub-score changes are more in line with each other.
3. Quantitative data can be calibrated with qualitative data to improve the precision of interpretations.
4. The dynamics of learning reveals that profoundly informed students may lower their total scores by shifting from narrow perspective right answers to broad perspective wrong answers as they read more into the test than it was intended to measure.
5. Using crosstabulations of two administrations converts analog behaviors (changes in thinking perspective) into digital form to enable statistical analysis.
6. Once this analysis has been completed, the student profiles can be converted back to their original analog structure in the form of the systematic changes of underlying cognition with time.

From Kuhn's[23] perspective, the increase in explained variability this alternative procedure provides (about three times as much) is sufficient to qualify this alternative as being more accurate and more effective than current practice.

These results agree with Heisenberg's conclusion. It depends upon how the observation procedures are undertaken for definitive interpretation. From these finding, at least some learning is a discontinuous, insight-based process that is not digital. The key to the research method reported here lies in the way that observations were converted to digital form so that they analog properties are not lost and can be restored to their original analog form, improving accuracy of interpretation.

"Is the universe digital of analog?" has an answer from this research, "It may be both." However, establishing this fact has requires a research method that captures the analog essence in digital form so that digital computations can be undertaken. When this is done, at least some of the analog properties can be restored for insightful (analog) interpretation.

References

Bond, T. G. & Fox, C. M. (2007). (2nd Ed.). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah, NJ: Lawrence Erlbaum.

Lord, Fredric. (1952). *A Theory of Test Scores.* Psychometric Monographs, Number 7. Philadelphia, PA. Ferguson.

Gorham, D. R. (1957). *Proverbs Test*. Missoula, MT: Psychological Test Specialists -- Item 18

Halloun, I. & Hestenes, D. L. (1985). Common sense about motion. *American Journal of Physics, 53*, (11) 1056 – 1065.

Heisenberg, W. (1927). The Uncertainty Principle. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik. 43*, (S) 172–198.

Kuhn, Thomas S. (1996). *The Structure of Scientific Revolutions* (3rd ed.). Chicago, IL: University of Chicago Press.

Luce, R. D. & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology, 1* (1), 1-27.

Piaget, J. (1985). *The Equilibration of Cognitive Structures: The Central Problem of Intellectual Development*. Chicago: University of Chicago Press. (New translation of The Development of Thought)

Papert, S. (1980) *Mindstorms: children, computers, and powerful ideas*. New York: Basic Books.

Powell, J. C. (2010a). *Making Peasants into Kings*. Bloomington, IN: Author House. ISBN: 978-1-4490-0634-1.

______ (2010b). Chapter 3: Testing as feedback to inform teaching. In J. Michael Spector, *et al.* (2010) *Learning and Instruction in the Digital Age: Making a Difference through Cognitive Approaches*. New York: Springer. ISBN: 978-1-4419-1551-1

_____ (2010c). Do profoundly informed students choose wrong answers, lowering their scores? Research report presented to the *Psychomertic Society*, Athens, GA

_____ (1977). The developmental sequence of cognition as revealed by wrong answers. *Alberta Journal of Educational Research XXIII*, (1) 43 – 51. ERIC EJ159335.

______ (1968). The interpretation of wrong answers from a multiple-choice test. *Educational and Psychological Measurement, 28*, 403-412.

_____, Bernauer, James and Agnihotri, Vishnu. (2010) An Analysis of Answer Selection Patterns from Multiple-Choice Items. This paper was presented to CELDA in Timisoara, Romania. ISBN: 978-972-8939-28-1

_____ & Clapp, Allan P. (2009). *Comparing RSAI with IRT and DIF: Using Live Data.* This is a PowerPoint® presentation given to the Psychometric Society, S. John's College, Cambridge, UK.

_____ & Shklov (1992) Obtaining information about learners' thinking strategies from wrong answers on multiple-choice tests. *Educational and Psychological Measurement, 52*, 847-865.

_____ & Miki, H. (1985). Answer anomalies, how serious? Nashville, TN: Paper presented to the *Psychometric Society*.

_____, Cottrell, D. J. and Lever, Margaret (1977) Schools I Would Like to See: An opinion survey with interesting possibilities. *Alberta Journal of Educational Research. XXIII* (3) 226 – 241.

Endnotes

[1] Heisenberg, W. (1927). The Uncertainty Principle. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik. 43*, (S) 172–198.

[2] Lord, Fredric. (1952). *A Theory of Test Scores.* Psychometric Monographs, Number 7. Philadelphia, PA. Ferguson.

[3] Powell, J. C. (2010a). *Making Peasants into Kings*. Bloomington, IN: Author House. ISBN: 978-1-4490-0634-1. (Pp. 204 , 206).

[4] *Source:* Powell, J. C., Bernauer, James and Agnihotri, Vishnu. (2010) An Analysis of Answer Selection Patterns from Multiple-Choice Items. This paper was presented to CELDA in Timisoara, Romania. ISBN: 978-972-8939-28-1

[5] Halloun, I. & Hestenes, D. L. (1985). Common sense about motion. *American Journal of Physics, 53*, (11) 1056 – 1065.

[6] A fuller account of this research is contained in:
Powell, J. C. (2010b). Chapter 3: Testing as feedback to inform teaching. In J. Michael Spector, *et al.* (2010) *Learning and Instruction in the Digital Age: Making a Difference through Cognitive Approaches*. New York: Springer.
ISBN: 978-1-4419-1551-1

[7] Gorham, D. R. (1957). *Proverbs Test*. Missoula, MT: Psychological Test Specialists -- Item 18

[8] Powell, J. C., Cottrell, D. J. and Lever, Margaret (1977) Schools I Would Like to See: An opinion survey with interesting possibilities. *Alberta Journal of Educational Research. XXIII* (3) 226 – 241.

[9] This formula is an adaptation of the *multinomial* procedure developed by Powell and Shklov (1992) Obtaining information about learners' thinking strategies from wrong answers on multiple-choice tests. *Educational and Psychological Measurement, 52*, 847-865 to bypass the linear dependency problem inherent in applying GLM solutions to test data.

[10] Powell, J. C. (1977). The developmental sequence of cognition as revealed by wrong answers. *Alberta Journal of Educational Research XXIII*, (1) 43 – 51. ERIC EJ159335.

[11] Piaget, J. (1985). *The Equilibration of Cognitive Structures: The Central Problem of Intellectual Development*. Chicago: University of Chicago Press. (New translation of The Development of Thought)

[12] Powell, J. C. & Miki, H. (1985). Answer anomalies, how serious? Nashville, TN: Paper presented to the *Psychometric Society*.

[13] Bond, T. G. & Fox, C. M. (2007). (2nd Ed.). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah, NJ: Lawrence Erlbaum. They suggest student laziness as an explanation (P. 22.)

[14] Powell, J. C. (1968). The interpretation of wrong answers from a multiple-choice test. *Educational and Psychological Measurement, 28*, 403-412. The details of this procedure are presented in this paper.

[15] Powell and Shklov (1992) *Op. cit.*

[16] Powell, J. C. (1968) *Op. cit.*

[17] Powell, J. C. (1977) *Op. cit.*

[18] Powell, J. C. and Clapp, Allan P. (2009). *Comparing RSAI with IRT and DIF: Using Live Data.* This is a PowerPoint® presentation given to the Psychometric Society, S. John's College, Cambridge, UK.

[19] Luce, R. D. & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology, 1* (1), 1-27.

[20] Powell, J. C. (2010c). Do profoundly informed students choose wrong answers, lowering their scores? Research report presented to the *Psychomertic Society*, Athens, GA

[21] Powell, J. C. (2010) *Op. cit.*

[22] Papert, S. (1980) *Mindstorms: children, computers, and powerful ideas*. New York: Basic Books.

[23] Kuhn, Thomas S. (1996). *The Structure of Scientific Revolutions* (3rd ed.). Chicago, IL: University of Chicago Press.